

# Dynamic magnetic resonance imaging: reliability of anatomical landmarks and reference lines used to assess pelvic organ prolapse

Suzan R. Broekhuis · Kirsten B. Kluivers ·  
Jan C. M. Hendriks · Mark E. Vierhout ·  
Jelle O. Barentsz · Jurgen J. Fütterer

Received: 5 August 2008 / Accepted: 21 October 2008 / Published online: 11 November 2008  
© The Author(s) 2008. This article is published with open access at Springerlink.com

**Abstract** The aim of this study was to determine the intra- and interobserver reliability of dynamic magnetic resonance (MR) staging in pelvic organ prolapse patients. In 30 patients with pelvic organ prolapse, dynamic MR images were assessed independently by two observers. Various anatomical landmarks to assess pelvic organ prolapse were used in relation to the pubococcygeal line, H-line, and mid-pubic line. Clinical measurement points were assessed in relation to the mid-pubic line. The intraclass correlation coefficients (ICC) were calculated to describe the intra- and interobserver reliability. Overall, the intra- and interobserver reliability of MR imaging measurements was excellent to good. The pubococcygeal line showed superior reliability (ICC range 0.70–0.99). The reliability of clinical measurement points, however, were only moderate (ICC range 0.20–0.96). The intra- and interobserver reliability of quantitative prolapse staging on dynamic MR imaging were good to excellent. The pubococcygeal line appears the most reliable to use.

**Keywords** Intraobserver reliability · Interobserver reliability · Magnetic resonance imaging · Pelvic floor · Pelvic organ prolapse · Reference line

## Introduction

Pelvic organ prolapse (POP) is a major health care problem, with 11% of women undergoing surgery for POP and/or urinary incontinence during life time, of whom 30% have repeat surgery [1]. Failure to identify all of the involved compartments may result in incomplete surgical repair with subsequent persistence or recurrence of the prolapse [2, 3].

Since the introduction of fast imaging sequences, magnetic resonance (MR) imaging has become a promising diagnostic tool in the assessment of the pelvic floor dysfunction. Many protocols concerning the MR imaging procedure and interpretation on POP have been introduced. Until now, only two studies have addressed the reliability of prolapse staging on dynamic MR imaging [4, 5]. However, the reference lines and anatomical landmarks which should be used to assess the presence and the quantitative staging of the prolapse is a subject of ongoing debate.

The aim of the present study was to determine the intra- and interobserver reliability of dynamic MR staging in POP patients.

## Materials and methods

Dynamic MR imaging of the female pelvic floor were reviewed in a cross-sectional observational study. MR imaging had been performed as part of routine clinical

---

S. R. Broekhuis (✉) · K. B. Kluivers · M. E. Vierhout  
791 Department of Obstetrics and Gynecology,  
Radboud University Nijmegen Medical Centre,  
P.O. Box 9101, 6500 HB Nijmegen, The Netherlands  
e-mail: S.Broekhuis@obgyn.umcn.nl

J. C. M. Hendriks  
Department of Epidemiology,  
Radboud University Nijmegen Medical Centre,  
Nijmegen, The Netherlands

J. O. Barentsz · J. J. Fütterer  
Department of Radiology,  
Radboud University Nijmegen Medical Centre,  
Nijmegen, The Netherlands

practice in patients with recurrent prolapse, especially in the posterior compartment, and in cases where the patient complaints did not correspond with clinical findings on request of the Department of Obstetrics and Gynecology at the Radboud University Nijmegen Medical Centre, which is a national tertiary referral centre for POP patients. The study was submitted and deemed exempt by the Institutional Review Board 1 August 2007.

A computer-generated list of 30 out of 63 patients who underwent dynamic MR imaging in the period of September 2005 through March 2007 were included. None of these images were regarded as unsuitable for assessment.

The MR datasets were assessed by two independent observers, an experienced radiologist (JF, 3 years of experience) and a novice observer (SB, half year of experience). One of the observers (SB) repeated the assessment of the set of images at least 1 month later, to determine the intraobserver reliability. Prior to the study, ten MR datasets not included in the study had been assessed in cooperation to reach consensus on interpretation of the MR images. A case record form with definitions of all measurement points and reference lines was used (see below). The observers were blinded to the clinical findings and the previous assessment of the images.

#### Imaging protocol

The dynamic MR imaging examination was performed with the patient in supine position with parallel and slightly flexed legs. Patients were requested not to void for 1–2 h prior to their examination. The rectum was opacified using 100–150 ml ultrasound gel. The urethra, bladder, and vagina were not opacified. No premedication was given. MR images were acquired using a 3T MR scanner (TIM TRIO, Siemens Medical, Germany) and an eight-channel body phased-array coil. MR images were obtained in the sagittal plane using a Half-Fourier acquisition single-shot turbo spin-echo sequence (2000 ms/90 ms repetition time/echo time; 150° flip angle). During the MR examination, the patient was asked to relax her pelvic floor muscles, to contract the muscles slowly, relax again, and then to increase the intra-abdominal pressure and strain in order to defecate. To assure that the patient followed the instruction given, all images were viewed online on the MR console. The MR examination time was 35 min. The images were analyzed offline at a later stage on a console with zoom facilities and electronic calipers. The midsagittal images were used to assess the prolapse.

#### Reference lines

The reference lines used to assess POP are shown in Fig. 1. According to literature, the pubococcygeal line was



**Fig. 1** MR image obtained at rest. Dynamic midsagittal Half-Fourier acquisition single-shot turbo spin-echo (2000/90; 150°) through the pelvis of a 62-year-old woman with symptoms of pelvic organ prolapse. The image shows the used reference lines. *PCL* pubococcygeal line, *H-line*, *MPL* mid-pubic line

defined as a straight line between the inferior rim of the pubic bone and the last visible coccygeal joint [5, 6], the H-line as a straight line between the inferior rim of the pubic bone and the posterior wall of the anal canal on the level of the impression of the puborectal sling [7], and the mid-pubic line as a line drawn through the longitudinal axis of the pubic bone, passing through its midequatorial point [8].

#### Anatomical landmarks and clinical measurement points

A whirl of urine in the bladder and/or a dent into the cranial portion of the bladder, seen during straining on the sagittal images, indicated adequate straining. The MR images during rest and maximal straining were assessed for POP with the use of the various anatomical landmarks in all three compartments in relation to the previously mentioned reference lines. Anatomical landmarks used for each compartment were the bladder base and bladder neck for the anterior compartment, the distal portion of the cervix or the vagina vault for the middle compartment, and the anorectal junction and the most antero-caudal point of the anterior rectal wall for the posterior compartment. The

distance from the anatomical landmarks to the different reference lines was measured perpendicular.

Besides the aforementioned anatomical landmarks, we have also introduced clinical measurement points on MR images to approximate point Ba, C, and Bp of the pelvic organ prolapse quantification (POP-Q) system, which refers to the most descended edge of the anterior vaginal wall, cervix/vaginal vault, and posterior vaginal wall, respectively [9]. For the anterior compartment, we used the most posterio-caudal point of the anterior vaginal wall, for the middle compartment the most distal point of the cervix or the vaginal vault, and for the posterior compartment the most antero-caudal point of the posterior vaginal wall. At rest, we additionally assessed the total vaginal length, measured from the fornix posterior or vaginal vault, following the contour of the vagina, until the crossing with the mid-pubic line. We assessed these measurement points in relation to the mid-pubic line because this line has been introduced by Singh et al. as a reflection of the hymenal remnants, which is the reference structure in the POP-Q system [8].

Qualitative staging is the most widely used method of prolapse staging. In addition to the quantitative staging, we have assessed the reliability of qualitative staging of prolapse. In case an individual measurement point descended below or above a reference line, the measurement was scored positive and negative, respectively.

Rectocele, enterocele, perineal descent, and genital hiatus

Furthermore, Fig. 2 shows the lines A and B, which are defined as an extended line of the anterior border of the anal canal [10–12] and the expected margin of the normal anterior rectal wall, respectively [13–15]. These lines were only applied in the presence of any outpouching of the anterior rectal wall. Additional measurements were then performed, which were the depth until the most antero-caudal point of the anterior rectal wall, and the area and perimeter of the outpouching in relation to lines A and B. The complete or incomplete evacuation of the outpouching during defecation was evaluated.

The presence of an enterocele was defined as any outpouching of the peritoneal sack, containing omentum and/or small bowel loops, into the rectovaginal space. The distance between the most distal point of the peritoneal sack and the vaginal vault, and the three reference lines, was measured, respectively.

Perineal descent was measured as a perpendicular distance between the pubococcygeal line and the anterior margin of the muscles sphincter ani [16, 17]. The dimension of the genital hiatus was defined as the distance between the inferior rim of the pubic bone and the posterior wall of the anal canal on the level of impression of the puborectal sling [7, 18].



**Fig. 2** MR image obtained at straining. Dynamic midsagittal Half-Fourier acquisition single-shot turbo spin-echo (2000/90; 150°) through the pelvis of a 58-year-old woman with symptoms of pelvic organ prolapse. The image shows the used reference lines A and B, applied in the presence of any outpouching of the anterior rectal wall

#### Statistical methods

The sample size calculation was performed before the start of this study based on the precision of the reliability. Twenty-three patients are needed to obtain a relative precision of 15% in the SE. In order to reach the number of 23 for all measurements of POP, a total of 30 patients are needed.

The intraobserver reliability was assessed of the first and second measurements of one of the observers, and the interobserver reliability was assessed of the first measurements of this observer and an additional measurement of the second observer. The intraclass correlation coefficient (ICC) was calculated to measure the reliability of the quantitative MR imaging measurements [19]. A linear mixed model was used to calculate the ICC of each specific measurement of POP, separately. A separate model was used to study the intraobserver ICC and for the interobserver ICC. The independent variable was the specific measurement of POP. The dependent random variable was “patient,” and the dependent fixed variable was observer one (first, second) in case of the intraobserver model and observer (one, two) in case of the interobserver model. The

**Table 1** ICC and mean difference, with 95% CI, of the within and between observer measurements of pelvic organ prolapse by anatomical landmarks and clinical measurement points in relation to three reference lines at rest and straining (n=30)

	Intraobserver		Interobserver	
	ICC (95% CI)	Mean difference in cm (95% CI)	ICC (95% CI)	Mean difference in cm (95% CI)
<b>Pubococcygeal line</b>				
<b>Rest</b>				
Bl	0.99 (0.98; 1.00)	-0.10 (-0.16; -0.05)	0.87 (0.74;0.93)	0.47 (0.22; 0.71)
Bn	0.98 (0.95; 0.99)	-0.03 (-0.11; 0.06)	0.92 (0.84; 0.96)	0.18 (0.04; 0.32)
C/V	0.95 (0.90; 0.98)	-0.08 (-0.26; 0.10)	0.83 (0.67; 0.91)	-0.15 (-0.46; 0.16)
ARJ	0.96 (0.91; 0.98)	-0.16 (-0.50; 0.17)	0.83 (0.67; 0.91)	0.54 ( 0.16; 0.93)
Rec	0.82 (0.66; 0.91)	0.19 (-0.10; 0.48)	0.70 (0.46; 0.85)	-0.27 (-0.66; 0.12)
<b>Straining</b>				
Bl	0.99 (0.97; 0.99)	-0.07 (-0.22; 0.09)	0.83 (0.68; 0.92)	-0.28 (-0.78; 0.22)
Bn	0.90 (0.81; 0.95)	0.11 (-0.12; 0.35)	0.70 (0.46; 0.84)	0.02 (-0.45; 0.50)
C/V	0.94 (0.88; 0.97)	-0.06 (-0.32; 0.21)	0.83 (0.66; 0.92)	-0.18 (-0.59; 0.23)
ARJ	0.84 (0.68; 0.92)	0.34 (0.00; 0.69)	0.84 (0.69; 0.92)	0.31 (-0.03; 0.64)
Rec	0.93 (0.87; 0.97)	0.17 ( 0.00; 0.35)	0.83 (0.67; 0.91)	-0.08 (-0.38; 0.21)
<b>H-line</b>				
<b>Rest</b>				
Bl	0.99 (0.97; 0.99)	-0.11 (-0.17; -0.06)	0.76 (0.55; 0.88)	0.48 (0.22; 0.73)
Bn	0.88 (0.77; 0.94)	-0.01 (-0.16; 0.15)	0.86 (0.73; 0.93)	0.23 ( 0.08; 0.39)
C/V	0.95 (0.90; 0.98)	-0.21 (-0.39; -0.03)	0.84 (0.69; 0.92)	-0.22 (-0.53; 0.08)
ARJ	0.91 (0.82; 0.96)	-0.15 (-0.34; 0.04)	0.73 (0.51; 0.86)	0.29 (-0.02; 0.59)
Rec	0.76 (0.56; 0.88)	0.01 (-0.36; 0.38)	0.73 (0.50; 0.86)	-0.31 (-0.73; 0.10)
<b>Straining</b>				
Bl	0.98 (0.95; 0.99)	-0.13 (-0.32; 0.07)	0.83 (0.67; 0.91)	-0.41 (-0.86; 0.04)
Bn	0.91 (0.82; 0.96)	0.07 (-0.17; 0.30)	0.69 (0.44; 0.84)	0.19 (-0.63; 0.25)
C/V	0.83 (0.67; 0.91)	-0.38 (-0.83; 0.06)	0.68 (0.43; 0.83)	-0.63 (-1.21; -0.04)
ARJ	0.75 (0.54; 0.87)	-0.13 (-0.58; 0.32)	0.45 (0.12; 0.70)	0.13 (-0.75; 0.50)
Rec	0.93 (0.86; 0.97)	0.02 (-0.17; 0.22)	0.86 (0.73; 0.93)	-0.18 (-0.47; 0.10)
<b>Mid-pubic line</b>				
<b>Rest</b>				
Bl	0.54 (0.23; 0.75)	0.12 (-0.11; 0.36)	0.50 (0.17; 0.73)	0.34 ( 0.07; 0.61)
Bn	0.94 (0.87; 0.97)	0.02 (-0.05; 0.09)	0.51 (0.19; 0.73)	0.34 ( 0.12; 0.55)
C/V	0.96 (0.92; 0.98)	-0.02 (-0.17; 0.14)	0.90 (0.79; 0.95)	-0.20 (-0.44; 0.03)
ARJ	0.96 (0.92; 0.98)	0.25 ( 0.11; 0.38)	0.77 (0.58; 0.89)	0.53 ( 0.23; 0.84)
Rec	0.84 (0.68; 0.92)	0.28 (-0.05; 0.62)	0.65 (0.38; 0.81)	-0.37 (-0.83; 0.10)
Tvl	0.79 (0.61; 0.90)	0.10 (-0.36; 0.55)	0.72 (0.49; 0.86)	-0.69 (-1.15; -0.22)
A	0.52 (0.20; 0.74)	0.28 (0.00; 0.56)	0.22 (-0.15; 0.53)	0.31 (-0.26; 0.87)
M	0.96 (0.91; 0.98)	-0.03 (-0.16; 0.16)	0.90 (0.79; 0.95)	-0.21 (-0.45; 0.02)
P	0.84 (0.68; 0.92)	0.30 ( 0.07; 0.52)	0.40 (0.01; 0.66)	0.19 (-0.28; 0.67)
<b>Straining</b>				
Bl	0.99 (0.98; 0.99)	0.23 ( 0.11; 0.34)	0.79 (0.61; 0.90)	-0.28 (-0.68; 0.12)
Bn	0.88 (0.75; 0.94)	0.29 ( 0.07; 0.52)	0.60 (0.31; 0.79)	-0.06 (-0.49; 0.37)
C/V	0.94 (0.87; 0.97)	-0.11 (-0.35; 0.13)	0.49 (0.16; 0.72)	0.32 (-0.48; 1.13)
ARJ	0.67 (0.41; 0.83)	0.75 (0.34; 1.17)	0.77 (0.57; 0.88)	0.68 (0.32; 1.04)
Rec	0.95 (0.89; 0.97)	0.35 ( 0.13; 0.58)	0.92 (0.84; 0.96)	0.17 (-0.11; 0.46)
A	0.94 (0.87; 0.97)	0.18 (-0.06; 0.43)	0.55 (0.25; 0.76)	-1.31 (-1.94; -0.69)
M	0.94 (0.87; 0.97)	-0.05 (-0.30; 0.20)	0.83 (0.68; 0.92)	-0.28 (-0.69; 0.12)
P	0.61 (0.33; 0.79)	0.36 (-0.20; 0.93)	0.45 (0.11; 0.70)	-1.54 (-2.27; -0.82)

*n* number of patients, *ICC* intraclass correlation coefficient, *CI* confidence interval, *Bl* bladder base, *Bn*=bladder neck, *C/V* most distal point of the cervix/vaginal vault, *ARJ* anorectal junction, *Rec* most antero-caudal point of the anterior rectal wall, *Tvl* total vaginal length, *A* most postero-caudal point of the anterior vaginal wall, *M* most distal point of the cervix/vaginal vault, *P*=most antero-caudal point of the posterior vaginal wall

mean difference between the two categories of the fixed variable and the ICC with their 95% confidence interval are presented. An ICC of more than 0.8 denotes excellent agreement, between 0.8 and 0.6 good agreement, between 0.6 and 0.4 moderate agreement, and below 0.4 poor agreement, respectively [20]. SPSS version 14.0 (SPSS, Chicago, IL, USA) was used to perform the statistical analysis.

## Results

Thirty women with POP were included. The median age was 52 years (range 32–76). The median parity was two (range 1–5). The median of the most descended point of prolapse during gynecological examination, according to the POP-Q system, was +1 cm (range minus 2–plus 7). Twenty-four women (80%) had a history of gynaecological surgery, consisting of surgery for pelvic organ prolapse ( $n=22$ ) and urinary incontinence ( $n=2$ ). Twenty patients had undergone previous hysterectomy, of whom two had a subtotal hysterectomy.

The results show systematic differences in the intra- and interobserver reliability in relation to the mid-pubic line, but also in relation to the pubococcygeal line and H-line. These differences are generally less than 0.5 cm which seems clinically irrelevant.

Table 1 shows the ICC and mean difference of the within and between observer measurements of pelvic organ prolapse at rest and straining by anatomical landmarks in relation to the pubococcygeal line, the H-line, and the mid-pubic line and by clinical measurement points in relation to the mid-pubic line.

The ICC of the anatomical landmarks were excellent to good, with some exceptions where inferior reliability came across. More specific, the interobserver ICC of the anorectal junction during straining (ICC=0.45) using the H-line and the intra- and interobserver ICC in the anterior compartment at

rest (ICC=0.54, 0.50 and 0.51, respectively) and the interobserver ICC during straining (ICC=0.60 and 0.49) using the mid-pubic line were moderate. The median interobserver ICC of the anatomical measurement points (Bl, Bn, C/V, ARJ and Rec) of the pubococcygeal line were the highest both at rest and during straining compared to the median interobserver ICC of the H-line and mid-pubic line (0.83 and 0.83 compared to 0.76, 0.69, 0.65 and 0.77, respectively). In conclusion, the interobserver agreement using the H-line and the mid-pubic line was somewhat disappointing. The median ICC of the pubococcygeal line at rest and during straining is substantially higher compared to the median ICC of the H-line and mid-pubic line, especially with regard to the interobserver reliability.

Table 2 displays the ICC and mean difference of the within and between observer measurements of any outpouching of the anterior rectal wall in relation to a straight line through the anterior border of the anal canal (line A) or in relation to the expected margin of the normal anterior rectal wall (line B). Both methods had a good to excellent intra- and interobserver reliability (ICC range 0.73–0.93).

The observers agreed on the presence of an enterocele in ten out of 30 patients. One more small enterocele has been assessed by one of the observers, which has been disregarded in the analysis. Table 3 presents the results of the ICC and mean difference of the within and between observer measurements of enteroceles during straining in relation to the pubococcygeal line, the H-line, and the mid-pubic line. The intraobserver reliability of the quantitative assessment for the three reference lines were excellent (ICC range 0.91–0.97), but the interobserver reliability for the pubococcygeal line and the H-line were only moderate (ICC=0.47 and 0.45, respectively).

Table 4 shows the ICC and mean difference of the within and between observer measurements of perineal descent in relation to the pubococcygeal line and the genital hiatus.

**Table 2** ICC and mean difference, with 95% CI, of the within and between observer measurements of any outpouching of the anterior rectal wall in relation to two different reference lines ( $n=30$ )

	Intraobserver		Interobserver	
	ICC (95% CI)	Mean difference in cm (95% CI)	ICC (95% CI)	Mean difference in cm (95% CI)
<b>Line A</b>				
Depth	0.80 (0.61; 0.90)	0.15 (−0.21; 0.50)	0.79 (0.61; 0.90)	0.58 (0.17; 0.98)
Area	0.88 (0.76; 0.94)	0.95 (−0.22; 2.12)	0.92 (0.81; 0.96)	1.04 (−0.08; 2.16)
Perimeter	0.73 (0.50; 0.86)	0.01 (−1.60; 1.63)	0.93 (0.86; 0.97)	0.19 (−0.69; 1.06)
<b>Line B</b>				
Depth	0.85 (0.72; 0.93)	−0.10 (−0.37; 0.16)	0.77 (0.57; 0.88)	0.69 (0.30; 1.07)
Area	0.92 (0.83; 0.96)	0.09 (−0.83; 1.00)	0.88 (0.76; 0.94)	0.74 (−0.47; 1.96)
Perimeter	0.75 (0.53; 0.87)	−0.65 (−2.37; 1.07)	0.91 (0.82; 0.96)	−1.13 (−2.12; −0.14)

$n$  number of patients, ICC intraclass correlation coefficient, CI confidence interval, Line A an extended line of the anterior border of the anal canal, Line B the expected margin of the normal anterior border of the anal canal

**Table 3** ICC and mean difference, with 95% CI, of the within and between observer measurements of enteroceles during straining in relation to three reference lines ( $n=10^a$ )

	Intraobserver		Interobserver	
	ICC (95%CI)	Mean difference in cm (95% CI)	ICC (95% CI)	Mean difference in cm (95% CI)
PCL	0.97 (0.89; 0.99)	-0.15 (-0.39; 0.09)	0.47 (-0.18; 0.84)	-0.06 (-1.21; 1.10)
H-line	0.91 (0.67; 0.98)	-0.45 (-0.95; 0.04)	0.45 (-0.21; 0.83)	0.09 (-1.32; 1.50)
MPL	0.95 (0.82; 0.99)	-0.22 (-0.62; 0.18)	0.78 (0.34; 0.94)	0.59 (-0.48; 1.66)

$n$  number of patients, ICC intraclass correlation coefficient, CI confidence interval, PCL pubococcygeal line, MPL mid-pubic line

<sup>a</sup>The observers agreed on the presence of an enterocele in ten out of 30 patients. One more small enterocele has been assessed by one of the observers, which has been disregarded in this analysis

Overall, the reliability of these measurements was good to excellent (ICC range 0.72–0.89), with the exception of the interobserver reliability at rest of perineal descent (ICC=0.52).

## Discussion

In the present study, we have determined the intra- and interobserver reliability of dynamic MR staging in POP patients, using the three most common reference lines at rest and during straining. The intra- and interobserver reliability of quantitative prolapse staging on dynamic MR imaging were generally excellent to good. However, systematic differences and inferior reliability were mainly seen in relation to the mid-pubic line.

To our knowledge, only two studies have questioned the reliability of dynamic MR imaging measurement of the pelvic floor. Morren et al. have assessed nulliparous female volunteers, using the pubococcygeal line as reference line and different anatomical reference points to evaluate the position of the pelvic organs [5]. Subsequently, Fauconnier et al. have assessed women with POP and used the mid-pubic line and a new-introduced reference line, the perineal line [4]. Both studies used the ICC to assess reliability. Morren et al. reported that most measurement error was due

to intraobserver reliability, which was not reproduced in this study. Fauconnier et al. have also reported excellent intra- and interobserver reliability of the MR imaging measurements.

Furthermore, Fauconnier et al. have introduced three measurement points on dynamic MR imaging similar to the clinical staging by the POP-Q system [4, 9]. In contrast with their findings, the reliability of these measurements was only moderate to poor on external validation in our study, with the exception of the middle compartment (i.e., the most descended point of the cervix or the vaginal vault).

In the present study, the cine-loop of images has been used, with very satisfactory reliability. Intra- and interobserver reliability can either be assessed in a cine-loop of images, where each rater picks the own image for assessment or in a single predefined image. Although, this difference may have an important influence on the results, the method used in the two previous studies on reliability has not been described.

The presence of POP on dynamic MR imaging has frequently been described whether or not pelvic organs descended below a certain reference line on Valsalva maneuver [12, 18, 21]. This is probably the most widely used method in clinical practice as well. In the present study, however, the reliability of this dichotomous distribution was overall poor or showed in virtue of expectation

**Table 4** ICC and mean difference, with 95% CI, of the within and between observer measurements of perineal descent in relation to the pubococcygeal line and the genital hiatus at rest and straining ( $n=30$ )

	Intraobserver		Interobserver	
	ICC (95% CI)	Mean difference in cm (95% CI)	ICC (95% CI)	Mean difference in cm (95% CI)
Perineal descent				
Rest	0.83 (0.68; 0.92)	0.08 (-0.14; 0.30)	0.52 (0.21; 0.74)	0.45 (-0.01; 0.91)
Straining	0.83 (0.67; 0.92)	0.17 (-0.10; 0.45)	0.72 (0.49; 0.86)	0.77 (0.36; 1.67)
Genital hiatus				
Rest	0.88 (0.77; 0.94)	-0.34 (-0.53; -0.15)	0.75 (0.54; 0.87)	0.39 (0.13; 0.64)
Straining	0.89 (0.78; 0.95)	-0.15 (-0.34; 0.03)	0.72 (0.49; 0.86)	0.30 (-0.06; 0.66)

$n$  number of patients, ICC intraclass correlation coefficient, CI confidence interval

a high agreement. These high agreements were due to the fact that all anatomical landmarks were at a far distance of the reference lines, and thus not discriminating. Consequently, a diagnostic test for POP, merely based on descend below a certain reference line, seems to be a nonvalid method for interpretation of dynamic MR images.

In literature, four different methods have been described to report on the presence or staging of an enterocele on MR imaging, i.e., (1) an enterocele was diagnosed in case of any outpouching of the peritoneal sack and its contents into the rectovaginal space [3, 7, 14, 15], (2) when the sack descended into the rectovaginal space for more than one third of the proximal vagina [12, 22–24], (3) when the sack descended more than halfway the vagina [25], or (4) when the peritoneal sack descended below a certain reference line [10, 11, 13, 26]. In the present study, the qualitative assessment (i.e., the first three methods) was excellent to good. The interobserver reliability of the quantitative assessment for the pubococcygeal line and the H-line, however, were moderate.

For the full validation of a certain method, data on, for example, validity and applicability needs to be assessed besides reliability. In the present study, reliability is the only parameter assessed. This limitation needs to be taken into account in the interpretation of the results.

Another issue, which needs to be considered, is the ease of use of reference lines. The pubococcygeal line is the most reliable to use, probably because this line is drawn between two fixed bony points. In addition, the pubococcygeal line is the most widely used reference line in the available literature, which is another advantage. Future studies are needed to establish whether the pubococcygeal line is also the preferable reference line with regard to validity, for example, compared to symptoms and the clinical staging of pelvic organ prolapse.

Dynamic MR imaging is mostly performed in the supine position, as open-magnet unit MR imaging, which are needed for prolapse assessment in the sitting position, are not very widespread. Bertschinger et al. has concluded in a study on closed-magnet unit dynamic MR imaging versus open-magnet unit dynamic MR imaging, i.e., supine vs. sitting position that the presence and severity grade of POP was concordant in the majority of patients [27]. Based on their conclusion, we think that our findings on reliability also apply when open-magnet unit dynamic MR imaging is used to assess POP.

The assessment of the MR images has been performed by an experienced radiologist (JF) and a novice observer (SB). Our data show that, in case of clear definitions of the reference line and measurement points, only a short training period was needed, to assess dynamic MR images regarding pelvic organ prolapse in a reliable manner.

In conclusion, the intra- and interobserver reliability of quantitative prolapse staging on dynamic MR imaging were generally good to excellent. The pubococcygeal line appears to be the most reliable to use, since the median ICCs are lower for the H-line and the mid-pubic line compared with the pubococcygeal line.

**Financial support** None.

**Conflicts of interest** None.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Olsen AL, Smith VJ, Bergstrom JO, Colling JC, Clark AL (1997) Epidemiology of surgically managed pelvic organ prolapse and urinary incontinence. *Obstet Gynecol* 89:501–506
- Kelvin FM, Maglinte DD, Hornback JA, Benson JT (1992) Pelvic prolapse: assessment with evacuation proctography (defecography). *Radiology* 184:547–551
- Rentsch M, Paetzel C, Lenhart M, Feuerbach S, Jauch KW, Furst A (2001) Dynamic magnetic resonance imaging defecography: a diagnostic alternative in the assessment of pelvic floor disorders in proctology. *Dis Colon Rectum* 44:999–1007
- Fauconnier A, Zareski E, Abichedid J, Bader G, Falissard B, Fritel X (2007) Dynamic magnetic resonance imaging for grading pelvic organ prolapse according to the international continence society classification: which line should be used. *Neurourol Urodyn* 27:191–197
- Morren GL, Balasingam AG, Wells JE, Hunter AM, Coates RH, Perry RE (2005) Triphasic MRI of pelvic organ descent: sources of measurement error. *Eur J Radiol* 54:276–283
- Lienemann A, Sprenger D, Janssen U, Grosch E, Pellengahr C, Anthuber C (2004) Assessment of pelvic organ descent by use of functional cine-MRI: which reference line should be used. *Neurourol Urodyn* 23:33–37
- Barbaric ZL, Marumoto AK, Raz S (2001) Magnetic resonance imaging of the perineum and pelvic floor. *Top Magn Reson Imaging* 12:83–92
- Singh K, Reid WM, Berger LA (2001) Assessment and grading of pelvic organ prolapse by use of dynamic magnetic resonance imaging. *Am J Obstet Gynecol* 185:71–77
- Bump RC, Mattiasson A, Bo K, Brubaker LP, DeLancey JO, Klarskov P, Shull BL, Smith AR (1996) The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *Am J Obstet Gynecol* 175:10–17
- Hetzer FH, Andreisek G, Tsagari C, Sahrbacher U, Weishaupt D (2006) MR defecography in patients with fecal incontinence: imaging findings and their effect on surgical management. *Radiology* 240:449–457
- Kelvin FM, Hale DS, Maglinte DD, Patten BJ, Benson JT (1999) Female pelvic organ prolapse: diagnostic contribution of dynamic cystoproctography and comparison with physical examination. *AJR Am J Roentgenol* 173:31–37
- Macura KJ (2006) Magnetic resonance imaging of pelvic floor defects in women. *Top Magn Reson Imaging* 17:417–426

13. Hubner M, Hetzer F, Weishaupt D, Hahnloser D, Clavien PA, Demartines N (2006) A prospective comparison between clinical outcome and open-configuration magnetic resonance defecography findings before and after surgery for symptomatic rectocele. *Colorectal Dis* 8:605–611
14. Mortelet KJ, Fairhurst J (2007) Dynamic MR defecography of the posterior compartment: indications, techniques and MRI features. *Eur J Radiol* 61:462–472
15. Roos JE, Weishaupt D, Wildermuth S, Willmann JK, Marincek B, Hilfiker PR (2002) Experience of 4 years with open MR defecography: pictorial review of anorectal anatomy and disease. *Radiographics* 22:817–832
16. Hoyte L, Schierlitz L, Zou K, Flesh G, Fielding JR (2001) Two- and 3-dimensional MRI comparison of levator ani structure, volume, and integrity in women with stress incontinence and prolapse. *Am J Obstet Gynecol* 185:11–19
17. Hsu Y, Summers A, Hussain HK, Guire KE, DeLancey JO (2006) Levator plate angle in women with pelvic organ prolapse compared to women with normal support using dynamic MR imaging. *Am J Obstet Gynecol* 194:1427–1433
18. Fielding JR (2002) Practical MR imaging of female pelvic floor weakness. *Radiographics* 22:295–304
19. Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86:420–428
20. Shrout PE (1998) Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 7:301–317
21. Pannu HK (2002) Magnetic resonance imaging of pelvic organ prolapse. *Abdom Imaging* 27:660–673
22. Gearhart SL, Pannu HK, Cundiff GW, Buller JL, Bluemke DA, Kaufman HS (2004) Perineal descent and levator ani hernia: a dynamic magnetic resonance imaging study. *Dis Colon Rectum* 47:1298–1304
23. Goh V, Halligan S, Kaplan G, Healy JC, Bartram CI (2000) Dynamic MR imaging of the pelvic floor in asymptomatic subjects. *AJR Am J Roentgenol* 174:661–666
24. Kaufman HS, Buller JL, Thompson JR, Pannu HK, DeMeester SL, Genadry RR, Bluemke DA, Jones B, Rychcik JL, Cundiff GW (2001) Dynamic pelvic magnetic resonance imaging and cystocolpoproctography alter surgical management of pelvic floor disorders. *Dis Colon Rectum* 44:1575–1583
25. Gousse AE, Barbaric ZL, Safir MH, Madjar S, Marumoto AK, Raz S (2000) Dynamic half Fourier acquisition, single shot turbo spin-echo magnetic resonance imaging for evaluating the female pelvis. *J Urol* 164:1606–1613
26. Tunn R, Paris S, Taupitz M, Hamm B, Fischer W (2000) MR imaging in posthysterectomy vaginal prolapse. *Int Urogynecol J Pelvic Floor Dysfunct* 11:87–92
27. Bertschinger KM, Hetzer FH, Roos JE, Treiber K, Marincek B, Hilfiker PR (2002) Dynamic MR imaging of the pelvic floor performed with patient sitting in an open-magnet unit versus with patient supine in a closed-magnet unit. *Radiology* 223:501–508